Face Recognition System with Personal Information Retrieval Using Deep Learning

A robust face recognition system combining deep learning-based facial feature extraction with structured personal information retrieval. Using FaceNet embeddings and Multi-Layer Perceptron classifier, the system accurately identifies individuals from images and augments predictions with metadata including age, gender, and nationality from a CSV dataset.

Presented By **Nandhakumar Chitra Ragupathy**

Introduction and Problem Statement

Face recognition is a critical task in computer vision with applications ranging from security and access control to personalized user experiences in entertainment and social media. Traditional methods rely on handcrafted features, which often lack robustness to variations in pose, illumination, and expression.

Modern approaches leverage deep learning, particularly convolutional neural networks and embedding-based representations, to achieve high recognition accuracy. This project implements a face recognition system using FaceNet embeddings to represent facial features in a high-dimensional vector space, classified using a Multi-Layer Perceptron to identify individuals.

Overall Accuracy

99.12% on 17-class celebrity dataset

Dataset Size

1,702 facial images for training and evaluation

System Architecture Overview

Image Input

Facial images captured from various sources

FaceNet Embeddings

128-dimensional vector representation extraction

MLP Classifier

Multi-layer neural network for identification

Information Retrieval

Personal details from CSV database

The system demonstrates the practical integration of computer vision and structured data for intelligent face recognition applications, supporting both single-image and batch-image inference with high confidence predictions.

Dataset and Preprocessing Pipeline

Dataset Description

The dataset consists of images of 17 celebrities, with multiple images per person. Examples include Angelina Jolie, Brad Pitt, and Leonardo DiCaprio. A total of 1,702 facial images were used for training and evaluation.

Face Encoding Extraction

Facial embeddings were extracted using the FaceNet library. The embeddings map each face image to a 128-dimensional vector, capturing essential distinguishing features. Images without detectable faces were discarded during preprocessing.

Label Encoding

Each individual's name was encoded numerically using LabelEncoder. The dataset was then split into training (80%) and testing (20%) sets, ensuring stratification to maintain class balance.

Tensor Conversion

The training and testing datasets were converted into PyTorch tensors for compatibility with the neural network training framework, enabling efficient GPU-accelerated computation.

Multi-Layer Perceptron Architecture

01

Input Layer

128 neurons corresponding to FaceNet embedding size

02

Hidden Layer 1

256 neurons with ReLU activation and Dropout (0.3)

03

Hidden Layer 2

128 neurons with ReLU activation and Dropout (0.3)

04

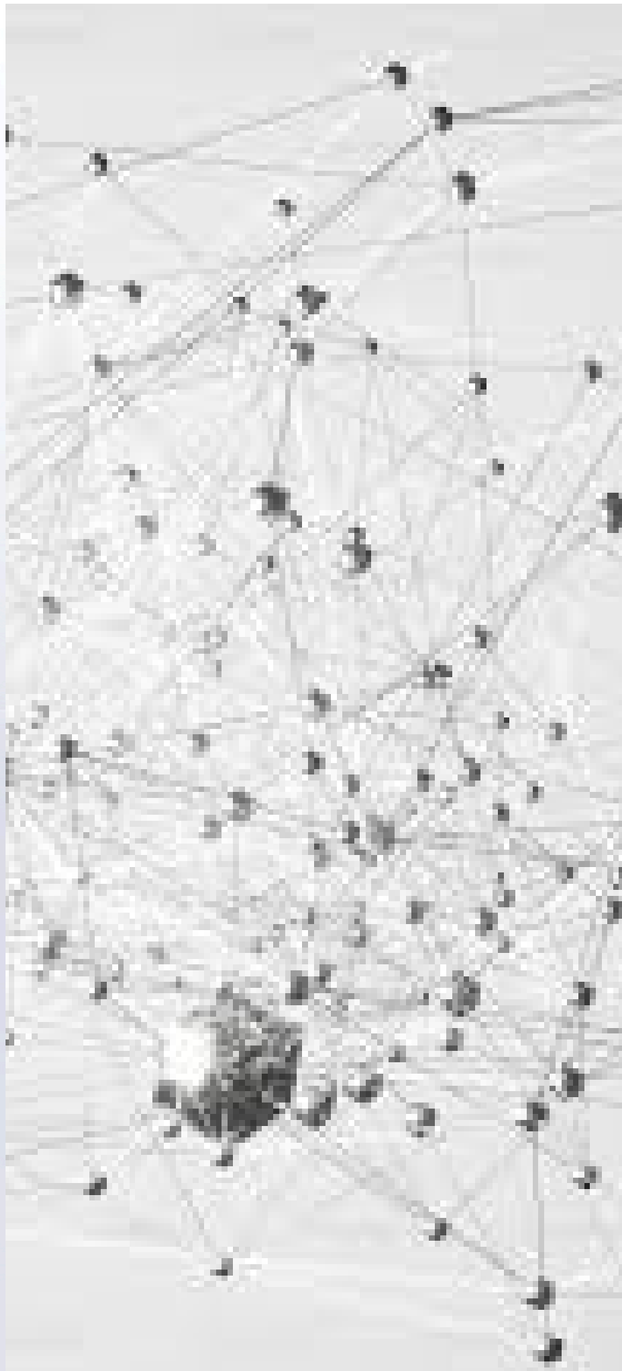
Output Layer

17 neurons representing classification categories

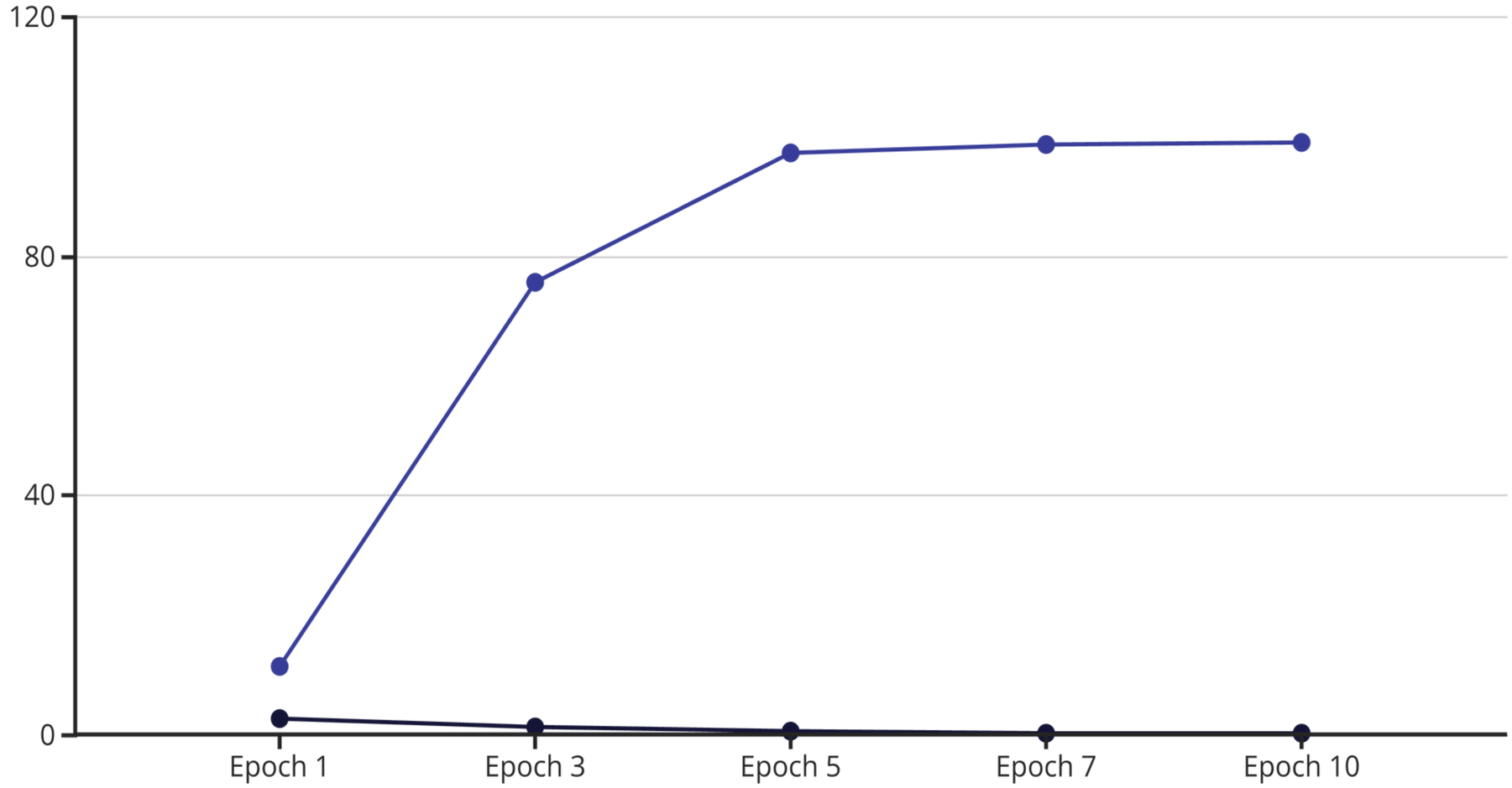
Training Configuration

- **Loss Function:** Cross-Entropy Loss
- **Optimizer:** Adam with learning rate 0.001
- **Epochs:** 10 training iterations
- **Batch Size:** 32 samples per batch

This architecture balances computational efficiency with the ability to capture complex patterns in the embedding space, achieving excellent convergence.



Training Performance and Convergence



The rapid reduction in loss and increase in validation accuracy indicate effective learning and minimal overfitting. The model achieved convergence quickly, demonstrating the quality of the FaceNet embeddings and the appropriateness of the MLP architecture for this classification task.

Evaluation Metrics and Performance

99.12%

Overall Accuracy

Classification performance on test dataset

1.0

Precision

Most classes achieved perfect precision

1.0

Recall

Near-perfect identification across categories

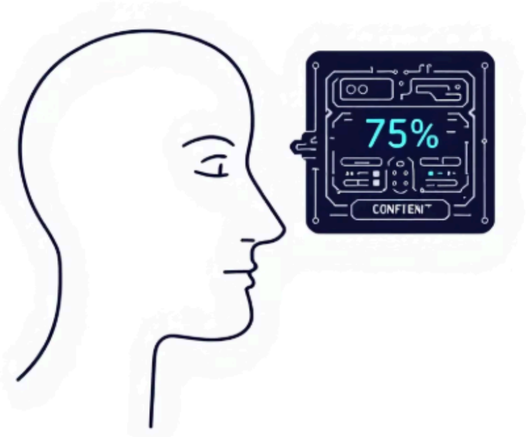
1.0

F1-Score

Balanced performance metric

Most classes achieved perfect scores of 1.0 across all metrics, with minor reductions for Megan Fox and Will Smith due to subtle visual similarity. The confusion matrix showed very few misclassifications, confirming robust performance across the entire dataset. This demonstrates the system's capability to distinguish between individuals with high confidence.

Prediction and Information Retrieval



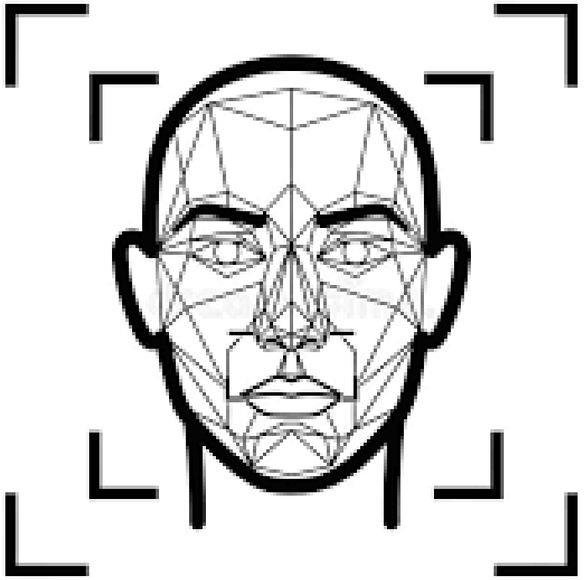
Single Image Prediction

The system processes individual images with high confidence:

> **anjel.jpg** → Predicted **Angelina Jolie** with 99.08% confidence

Batch Image Processing

All test images were correctly classified with high confidence ranging from 93% to 100%, demonstrating consistent performance across the entire evaluation set.



Angelina Jolie

Age: 53 | **Gender:** Female | **Nationality:** American

Brad Pitt

Age: 45 | **Gender:** Male | **Nationality:** American

Leonardo DiCaprio

Age: 48 | **Gender:** Male | **Nationality:** American

Predicted names are cross-referenced with a CSV file containing personal attributes, demonstrating the system's capability to provide enriched predictions beyond simple recognition.

Discussion and Key Findings

High Accuracy

Combining FaceNet embeddings with MLP classifier yields 99.12% accuracy, demonstrating robust performance across multiple images per individual

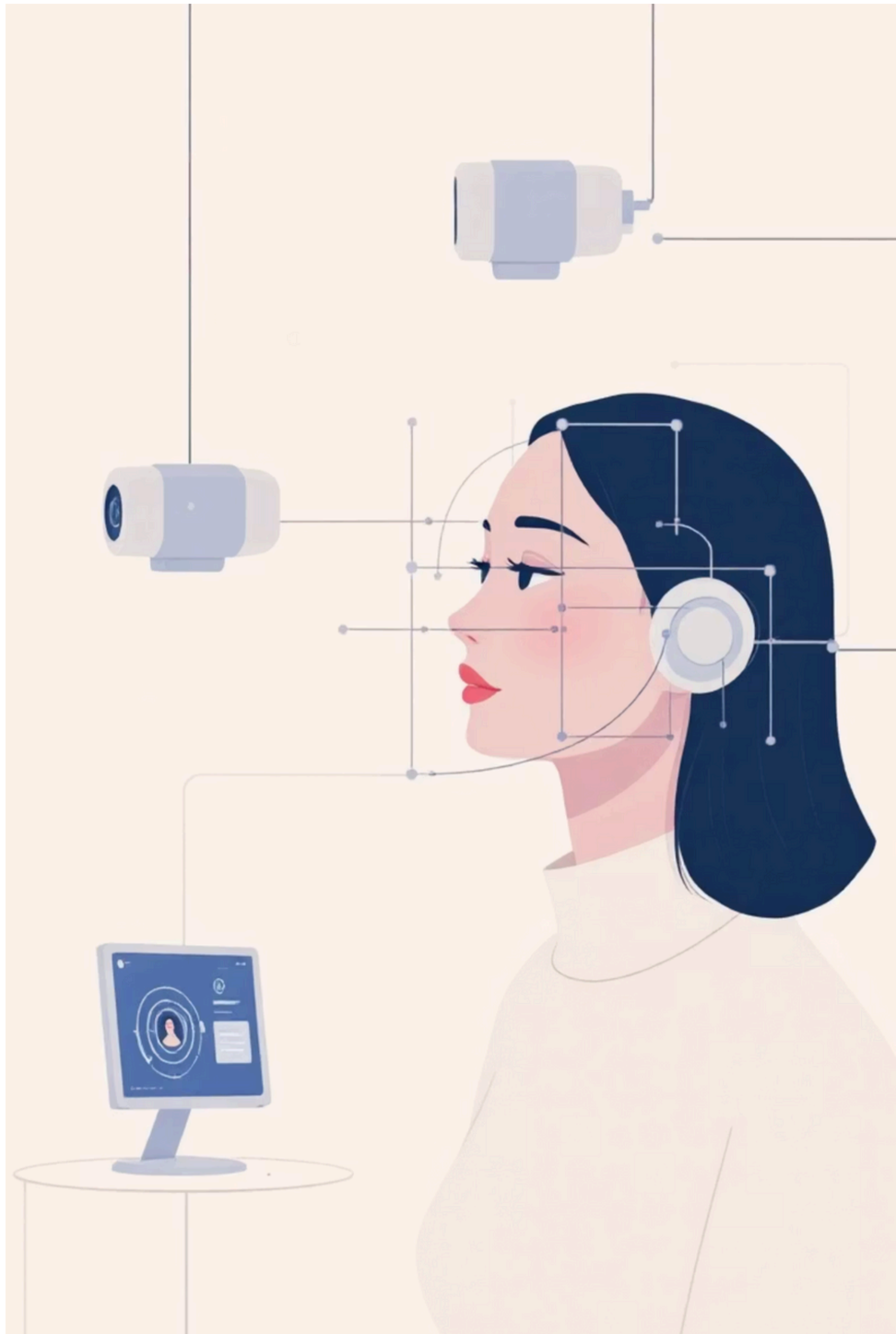
Dual Capability

System supports both single and batch inference modes, providing flexibility for different application scenarios

Contextual Information

Integration with CSV-based personal information database provides enriched predictions suitable for real-world applications

Minor misclassifications occurred for individuals with similar facial features, suggesting that increasing dataset diversity or using more complex neural network architectures such as CNNs or fine-tuned embeddings could further improve accuracy. The modular design enables straightforward enhancements and scaling.



Conclusion and Future Directions

Project Summary

The project successfully develops a face recognition system with personal information retrieval, achieving high accuracy and reliable performance. The combination of deep learning-based embeddings and structured data retrieval demonstrates the practical integration of computer vision and information systems.

The modular and scalable design allows for future enhancements, including real-time recognition, larger datasets, and additional personal attributes. This work underscores the potential of AI-driven solutions for intelligent face recognition applications.

Expand Dataset

Include more individuals and diverse images for improved generalization

Real-Time Recognition

Incorporate video stream processing for live identification

Advanced Architectures

Explore CNNs or fine-tuned FaceNet embeddings for enhanced accuracy

Rich Attributes

Add more personal attributes for comprehensive information retrieval

Application Deployment

Deploy as web or mobile application for practical use cases